DOCUMENT RESUME

ED 295 974 TM 011 781

AUTHOR Magliaro, Susan G.; Burton, John K.
TITLE The Cue Selection and Hypothesis Generation by Expert Reading Diagnosticians.
PUB DATE Apr 88
NOTE 45p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cues; Elementary Education; *Hypothesis Testing; Preadolescents; Reading Consultants; *Reading Diagnosis; Reading Difficulties
IDENTIFIERS *Cue Selection; Diagnostic Skills; Experts

ABSTRACT

The cue utilization and hypothesis testing strategies of expert reading diagnosticians were examined throughout the diagnostic process. Diagnostic processes of three diagnosticians (a reading specialist, a learning disabilities specialist, and a classroom teacher), who were very different in background and teaching assignment, were studied at work with three girls and three boys (aged 8-11 years). The major data sources included audiotaped and videotaped verbal reports from each diagnostic session, diagnosticians' notes made during the preparation and interaction stages, and the written diagnoses. Selected cues and hypotheses were generally consistent within diagnosticians. These diagnosticians were not as inconsistent as the research of J. F. Vinsonhaler et. al. (1983) indicated. In the entire process, diagnoses and prescriptions were largely consistent, although the diagnosticians may have differed in the process or the labels given the observed reading difficulties. From the examination of cues to the generating and testing of hypotheses, the process tended to arrive at similar results. Eleven figures and two tables summarize data from the experimental sessions. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED295974

THE CUE SELECTION AND HYPOTHESIS GENERATION BY EXPERT READING DIAGNOSTICIANS

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
SUSAN G. MAGLIARO
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Susan G. Magliaro

and

John K. Burton

Virginia Tech

A paper presented at the annual meeting of the
American Educational Research Association
April, 1988

TM011781

BEST COPY AVAILABLE

THE CUE SELECTION AND HYPOTHESIS GENERATION BY EXPERT READING DIAGNOSTICIANS

Perspective

In the area of cognitive science, what experts know about their work environment and how they use expert knowledge to identify problems have received increasing attention in recent years (Berliner, 1986; Bloom, 1986; Leinhardt & Greeno, 1986; Schon, 1983). Since DeGroot's (1966) initial study of chess masters, the cognitive orientation toward the study of expertise has maintained that qualitative as well as quantitative differences exist between the performance of experts and novices. Such differences have been found for experts and novices' declarative knowledge and procedural knowledge, as well as for the interaction among these knowledge structures. Issues such as these have become the focus of investigations of human problem solving processes in such domains as computer programming (e.g., Adelson, 1981), chess (e.g., Chase & Simon, 1973), physics (e.g., Chi, Feltovich & Glaser, 1981), and education (e.g., Housner & Griffey, 1985).

Problem solving is described as a complex thinking process involving goal-directed activity across a sequence of stages (Polya, 1957). Polya (1957, 1968) outlines four general stages of the process: (a) understanding the problem, (b) devising a plan, (c) carrying out the plan, and (d) looking back. Of these four stages, the first -- understanding the problem -- is arguably the most critical for a successful solution (Newell & Simon, 1972). The solver must gather information about the problem, find out what is known and unknown, and assess the

conditions in and under which the solution must be generated. Consequently, success at this stage requires the solver to be knowledgeable of the content area and identify the relevant cues that will lead to an accurate identification of the problem (Neves & Anderson, 1981).

Cue relevance, and subsequent cue effectiveness, is based on the principle of encoding specificity (Tulving & Thomson, 1973), and directly related to how information is initially learned. "Specific encoding operations performed on what is perceived determine what is stored, and what is stored determines what retrieval cues are effective in providing access to what is stored" (p. 369). Since cue effectiveness depends upon the stimulus conditions present at the time of attempted retrieval, which are likely to vary somewhat from the original encoding context, cues appearing in different contexts may access different bodies of information. The strength of the association between the cues and stored knowledge, developed through hours of reinforced practice, becomes a key factor. That is, cues that prove to be most effective in activating stored knowledge are naturally and strongly associated with the content of that knowledge base.

As the task environment becomes more complex and the number of competing cues increases, strategies for cue selection must be considered (Bruner et al., 1956). Strategies are implemented to obtain information relevant to the goals of the inquiry, reduce the amount of cognitive effort, and/or reduce the risk of making an error. Strategies will vary on each of these dimensions and

what is eventually implemented will be based on individual differences and contextual requirements.

Basic to all strategies is the valence or marker of the available cue (Bruner, et al., 1956). That is, does the cue represent a positive or negative instance of the concept? For example, in identifying a canary, singing and yellow would carry positive valences, while a bill and talons would be marked with negative valences. Associated with the dichotomous cue judgment (i.e., positive versus negative) is the much "grayer" judgments made about the relative importance or salience of a particular cue under a particular set of circumstances (Trabasso, 1963). Undergirding both the notion of cue valence and salience is the amount of prior experience or training the individual has had in the particular domain. Cues are attended to and concepts identified according the set of learned "rules" the individual has developed through practice.

These rules, however, are not restricted to single cues judged in a linear, one-at-a time sequence (Seibel, 1963). When situations become complex (as would be found in naturally occurring contexts), people impose decision rules on the context to integrate or "chunk" the noticed information into multiple cue judgments, thereby decreasing the cognitive load. These decision rules or judgments often reflect the individual's desired rules and policies somewhat imperfectly, due to systematic biases and random error (e.g., Hammond & Summers, 1972; Slovic & MacPhillamy, 1974). Most systematic biases, or perceptions about cue importance, will have been generated through the individual's

prior experience or training in the task in the context (Trabasso, 1963). Consequently, consistent implementation of individual policy appears to be related to the degree of cognitive skill for the context and the task.

Epitomizing problem identification, or the understanding of the problem, is the process of diagnosis. Pioneering work in this area has been conducted by Elstein and colleagues (e.g., Elstein, Kagan, Shulman, Jason & Loupe, 1972; Elstein, Shulman, Sprafka, 1978; Shulman & Elstein, 1978) using medical diagnosticians -- physicians. Their findings indicate that experienced physicians appeared to leap directly to a small array of provisional hypotheses very early in their meetings with patients. Further, these provisional hypotheses seemed to be generated out of the physicians' background knowledge of medicine, including their range of specific experiences, associated with the problematic cues recognized in the early stages of interaction with the patients.

Recent attempts to extend such work into the area of reading diagnosis by Vinsonhaler and colleagues (e.g., Gil, Hoffmeyer, VanRoekel, Vinsonhaler & Weinshank, 1979; Lee & Weinshank, 1978; Vinsonhaler, Weinshank, Wagner & Polin, 1983) has indicated that, unlike medical diagnosticians, reading diagnosticians evidence little, if any, intra- or interdiagnostician consistency (0.20, 0.10, respectively). Furthermore, their prescriptions appear to be unrelated to their diagnoses. A close examination of this research reveals that Vinsonhaler and his colleagues wanted to maintain a controlled, experimental environment, and consequently

restricted the availability of extraneous cues. This goal focused the studies away from the more "natural" settings which had succeeded in medicine (Elstein et al. used "well-trained" actors in actual examination room settings) toward more discrete events and measures typified by the representations of "clients" by boxes of data. Individual cues (e.g., a test score) could be accessed only one-at-a-time in a sequential manner. This method makes it virtually impossible to discern interactions between multiple cue judgments, pattern recognition and knowledge structures. In addition, these studies ignored the powerful effect of context (i.e., the availability of effective cues) on domain-specific, skilled memory.

Related to the limitations of Vinsonhaler's data collection procedures is the resultant limitations of the data analysis. "The major data for analysis was the set of statements in the written diagnosis" (p. 143, Vinsonhaler et al., 1983). While this technique may have resulted in the major evidence that supported the diagnostic statements, the cues and hypotheses that may have been used to confirm or disconfirm evidence for the final statement were eliminated.

With this in mind, our investigation was designed to replicate the work of Vinsonhaler and to extend the study of reading diagnosis into more ecologically valid settings. Specifically, this study examined the cue utilization and hypothesis testing strategies of expert reading diagnosticians throughout the entire diagnostic process. Settings ranged from "boxed" client files to work with live clients. The specific

questions which guided this research were:

1. What are the number and nature of cues selected by each diagnostician across the experimental sessions?

2. What are the number and nature of hypotheses generated by the selected cues?

3. How consistent are the cue selection and hypothesis/test strategies within and across diagnosticians?

4. How consistent are the diagnoses across diagnosticians?

5. Are there differences in cue selection and hypothesis/test strategies in live interaction versus videotaped versus "boxed" client sessions within and across diagnosticians?

6. Are there differences in diagnoses in live interaction versus videotaped versus "boxed" client sessions within and across diagnosticians?

As the data base is extremely large, with many dimensions of contrast, the purpose of this paper is to examine the diagnostic process as it occurs in naturalistic settings and to examine the consistency among diagnosticians who work with the same client. To that end, this document describes the diagnostic processes of three diagnosticians who were judged to be the most different in terms of training, present teaching assignment, beliefs about reading, and goals for diagnosis. We will focus on two of the five experimental sessions: the interaction with an actual client, to illustrate the typical, yet complex, diagnostic process; and the session in which they all "interact" with the same client via a videotape, to examine interdiagnostician consistency. The following section outlines the entire method in order to set the context for our research. Aspects of the method that are particularly relevant to this paper are highlighted.

Method

For the entire investigation, the participants were three reading specialists, a learning disabilities specialist, and a classroom teacher. These five diagnosticians were selected from a pool of participants who were initially recommended by their public school supervisors and/or university professors. In order to ensure a diverse sample, final selection was based on differences in professional training and conceptions of reading as determined by survey (Leu & Kinzer, 1987). Each person in the final purposive sample has a master's degree in either reading or learning disabilities and at least seven years of teaching experience (range = 7 to 36 years). Our focal diagnosticians include one of the reading diagnosticians, the LD specialist, and the classroom teacher. See Figure 1 for a description of their backgrounds, typical diagnostic procedures, and assessments.

The clients (ages 8-11, 3 girls and 3 boys) with whom the diagnosticians worked were enrolled in a university reading clinic program. All clients had a history of reading problems in their public school experience and were recommended to the clinic program by their parents. All of the experimental sessions in which the diagnostician worked with the actual client were held in the university clinic, in rooms that were familiar to the clients.

Five experimental sessions were conducted with each of the five participants, yielding a total of 25 sessions (see Figure 2). Each session consisted of a preparation stage (i.e., examining the client's file), an interaction stage (i.e., working

with that client in either live, videotaped, or "boxed"/audiotaped format), a written diagnosis stage, and a "wrap-up" stage that elicited a diagnostic summary statement, prescription, and reflection on each particular session. (A master list of the contents of clients' case information appears in Figure 3.) The tasks were varied by the mode of interaction with the client, the amount of time allocated to each stage (based on expert judgment), the nature of the verbalization (descriptive or explanatory), and whether a think-aloud or stimulated recall was employed. (Please see Ericsson & Simon, 1984, for a discussion of guidelines for using verbal reports, as well as the varying degrees of insight and nature of reponses and possible distortion inherent in this research methodology.)

As shown in Figure 4, five of the clients were rotated through experimental sessions I-IV, and all diagnosticians worked with the sixth client in experimental session V as a control condition. (This is Peter in Figure 4.) The focal clients are highlighted in either a circle or a box. The task order was counterbalanced within the constraints of parents' schedules and inclement weather (see Figure 5). Again, focal sessions are highlighted in the figure. All preparation and interaction stages were audio and videotaped, all stimulated recall and wrap-up stages were audiotaped.

The major data sources were the verbal reports that were audio and videotaped from each diagnostic session, the diagnosticians' notes made during the preparation and interaction stages, and the written diagnoses. The unit of analysis was the

experimental session. The total number and nature of single cues, multiple cues, hypotheses, diagnoses, and prescriptions were examined and compared within and across diagnostician and client. Two additional variables were examined as a result of their apparent importance and relevance to the final diagnoses. These were missing cues (i.e., information requested by the diagnostician, but not available) and inferred cues (inserted or inferred information for missing or unavailable cues, based on best available evidence). A qualitative analysis (Spradley, 1980) of the preceding variables was conducted. Each single, multiple, missing, and inferred cue and hypothesis was labeled in terms of its meaning and organized into a taxonomy. Diagnoses and prescriptions were compared to the selected cues within the session, within clients across sessions, and across diagnosticians across sessions. To reiterate, the purpose of these extensive comparisons was to construct an ecologically valid picture of the diagnostic process.

Results and Discussion

Our findings are presented on two levels. On a general level, a brief examination of five overall findings are reported to address the specific research questions and to set the context for the second level of results. For the more in-depth analysis, we next offer a close look at three of our participants and the results emanating from what we (and all five of the diagnosticians) considered to be an ecologically valid picture of their typical diagnostic procedures (i.e., experimental session

I, interaction with three of the actual clients), and then their inter-diagnostician consistency via the control condition (i.e., experimental session V, videotaped interaction with the same client across diagnosticians). General Results

Single cues. The findings indicate first that, across all conditions and diagnosticians, the mean number of different single cues considered for each experimental session was 293. Six domains of single cues emerged from the data: general/family background, health/medical information, school information, general intellectual ability, level of achievement, general interactions or events during the testing situation (see Figure 6). For all domains except the last (related to the testing situation), cues were gleaned from information in the records and from the interaction through questioning and observation. In those cases where access to the client was most limited (the "boxed" client sessions), the diagnosticians extracted appreciably more information from the files. In the videotaped conditions, their major focus was on the interaction they were viewing, with equal or less emphasis on the file. In the live conditions, the interaction served to corroborate hypotheses or questions formulated from the file information, as well as allow the diagnosticians to complete their own agendas.

As a general data summary, the domains of types of cues emanated from four major sources. These sources are:

1. the file information, which included family background, health information, grades, standardized test results;

2. behavioral observations during the interaction, which included the nature and level of physical activity, oral reading, question/answer behavior;

3. the task environment, that is the tests, the testing situation, the interaction between the client, diagnostician, tests, and testing situation; and

4. patterns of cues or multiple cue judgments that were combinations of cues from any single or combination of the general cue types.

Multiple cues. Secondly, multiple cues could not be organized into a taxonomy due to their interactive nature across the various sources of information. In addition, as the diagnostic process was traced through the stages, the multiple cues built on each other and "grew" to form related chunks of information used by the diagnosticians in identifying the problem. For example, in diagnosing Brett, Amy built a picture of his general health. During the preparation stage, she began with grouping all information from the health file about vision, hearing, and speech separately. She then examined the kindergarten checklist that reported information about the client's visual and auditory discrimination, and oral expression. By the time of stimulated recall, Amy had grouped the health information on vision and hearing together, then hearing with speech, and compared it with the information found on the checklist. Finally she grouped the fact that no deleterious childhood diseases were experienced with the total hearing and vision chunk (health plus checklist information) to negate any possibility of the client's health impacting on his ability to read. By the time Amy gave her final summary, she had drawn the conclusion that client A had no problems with health related to reading. Occurring in every experimental session with every diagnostician, this process appears quite similar to the chunking

of information demonstrated by experts in the studies cited in the introduction of this paper (e.g., Adelson, 1981; Chase & Simon, 1973; Leinhardt & Greeno, 1986).

Missing cues and inferences. A third aspect of the data surfaced as we began the analysis process. We found the number and nature of missing cues and resultant inferences emerged as a critical concern of the diagnosticians as they attempted to identify the clients' problems. As stated earlier, missing cues were defined as specific bits of information requested but not available. Inferences were defined as specific bits of information instantiated by the diagnostician based on what they considered to be their best guess. For example, given the grade level and book level, one diagnostician inferred that the teacher does not group her students for reading.

The number of these cues fluctuated, and seemed to depend upon two factors: the experimental session and the consistency between different pieces of information. Fewer missing cues were cited in sessions with live clients, than videotaped sessions, than "boxed" client sessions. Categories of missing cues paralleled the categories of single cues (see Figure 6). Those missing cues that were cited most often related to information the diagnosticians would extract from their own typical assessments, information from the child regarding such things as bedtime, attitudes toward school, classmates, teachers, and reading, and favorite games, hobbies or interests, and information from the teacher about the client's in-class behavior or performance. One diagnostician complained that she felt

"handstrung" without the ability to talk with the client's teachers and observe the client in the classroom setting.

Hypotheses. The fourth general observation was that the average number of hypotheses considered per experimental session was 14. (See Figure 7 for the taxonomy of the categories of tentative hypotheses cited by the diagnosticians.) It should be noted that this number includes "mini"-hypotheses or tentative conclusions about certain cues or multiple cues that were used to build a case for the final diagnosis. For example, to reach a diagnosis that a client had poor word attack skills, a diagnostician would initially gather information about the client's knowledge of consonants, consonant blends, and vowels using real and nonsense word assessments. Then, additional information would be gathered to ascertain the client's knowledge about syllables, prefixes, and suffixes. This body of information would inform the diagnostician about the client's word attack skills, and would subsequently be compared against the information the diagnostician gathered on the client's knowledge of sight words and word attack skills demonstrated in reading passages. Throughout this process, the diagnostician would verbalize tentative hypothetical statements regarding phonetic knowledge that she would attempt to confirm or disconfirm (e.g., "This client seems to have a problem with medial vowels.").

Consistency within and across diagnosticians. A fifth general finding was that, unlike the studies by Vinsonhaler and his colleagues (1983), the selected cues and hypotheses were

consistent within diagnosticians. Any variation within diagnosticians was due to the cues or time that was available in the different procedures, or to the individual client. Like the Vinsonhaler studies, however, we did find that the inter-diagnostician analyses revealed lower consistency, but that these process differences did not result in marked differences in the final diagnoses or prescriptions. That is, the diagnosticians may have differed in the process or the labels give to observed reading difficulties, but they reached somewhat similar conclusions about diagnoses and prescription for a specific client. While Vinsonhaler reported his diagnostician inconsistency to be unrelated to training experiences, we found that the training experiences were often cited as reasons for cue selection interpretation. In addition, the diagnosticians' present instructional assignments and their conceptions of reading were two other factors that appeared to account for inter-diagnostician inconsistencies.

The most interesting point of consistency across diagnosticians was related to the availability of cues. That is, in order to arrive at more accurate diagnoses and to conduct their typical diagnostic activities, in 16 of the 25 sessions the diagnosticians wanted more time to interact with the clients over a series of sessions. One diagnostician reported that "the most important information was working with her (the client), in the teaching situation." Further, the diagnosticians wanted to consult with other teachers, specialists, and the client's parents to gain a more complete picture of the clients as

readers. In five of the 25 sessions where critical cues were missing, the diagnosticians refused to make a diagnoses. In all cases the prescriptions or suggestions for further testing and instruction were directly related to the tentative hypotheses and diagnostic statements. A Closer Look at the Results of Two Experimental Sessions

As a way of illustrating the richness of the "pictures" of clients as readers that the diagnosticians attempted to develop, we will look at the diagnostic problem solving (in terms of the variables outlined in the general findings) of three diagnosticians, each coming from a different perspective. (Please refer back to Figure 1 for a brief outline of the background experiences and diagnostic procedures and assessments for the three target diagnosticians.) First, the findings from their interaction with live clients are reported. Second, the findings from the control condition, in which they all were able to "interact" with the same client in the same way, are presented. These two excerpts from the broader study were chosen to illustrate what diagnosticians typically do and consider as they attempt to identify reading problems, and to re-examine the findings by Vinsonhaler regarding the apparent inconsistency within and across diagnosticians.

Working with the real thing... Not knowing anything about the clients other than the children would be between the ages of 8 and 11 and had been enrolled in the university reading clinic, the diagnosticians came to the "live" sessions with their own totebags full of assessments. For each diagnostician, these

assessments consisted of those items outlined in Figure 1. Ann had everything she needed. Judy brought those standardized assessments that were not already in the clinic. Sarah had spent time investigating the library shelves to make sure she would be able to pull books for an informal reading sample.

For this first session, the preparation stage was limited to 30 minutes, they could take less time if desired. Judy and Sarah conducted a selective, purposive search through their clients' folders, focusing on those cues outlined in Figure 8. Judy, the LD specialist, spent most of her time with the standardized tests typically used to diagnose learning disabilities, while Susan, the classroom teacher, examined report card grades, teacher comments, and the curriculum referenced tests of mastery. Amy attempted to skim the entire folder, much the same as she does in the beginning of the school year when trying to identify potential students for her remedial reading class.

When it came time to choose assessments, Judy chose on the basis of the data she had collected on her client regarding a possible language disability, as well as a standardized reading assessment. Amy decided to go with her basal IRI, and some informal assessments on phonics, as her client seemed to have problems in that area. Sarah wasn't able to get a clear picture of her reader, as there may have been some emotional problems impacting on the reading difficulty. She decided to use the Ekwall IRI after her client had a chance to read some pages from books in the library that the client would choose.

The diagnosticians had up to one hour to meet with their

respective clients, after which they were allowed to go back into the file for additional information. Sarah did not ask for any other information, Amy just wanted to look a little more closely at some of the information she had, and Judy asked for the tutor's report from the clinic.

Table 1 reports the number of single cues, multiple cues, missing cues and inferences, and hypotheses verbalized by the three diagnosticians over the entire experimental session. Over the entire session, all three examined approximately the same number of different single cues, with Sarah verbalizing about 20% more total single cues than the other two. For all three, most of the single cues were extracted from the preparation stage, from their written notes, and during the stimulated recall stage. For Sarah and Amy especially, the stimulated recall stage represented a time to talk about many cues not mentioned previously.

The taxonomy of single cues derived from these sessions is found in Figure 6. Each of these three diagnosticians sampled cues from every major category. Emphasis on any one category depended upon the salient information extracted from the particular client's file and the diagnostician-client interaction. In her effort to ensure Tonita's security with the testing situation, Sarah focused more on the cues emanating from their interaction, with secondary emphasis the reading performance. Judy focused on Matt's responses, but was very sensitive to what she perceived his not wanting to be in the testing situation. Amy was able to study Brett's reading

performance, as he was very friendly and extremely cooperative.

The patterns of multiple cue selection were similiar across diagnosticians, with Amy citing somewhat fewer multiple cues during the preparation stage. The number of multiple cues generally increased toward the end of the experimental session as the diagnosticians were putting more pieces together to develop a picture of the child as a reader. As mentioned earlier, as the diagnosticians became more familiar with each case, and single bits of information began to tie together to "make sense," the size of the multiple cue judgments (i.e., number of bits comprising the whole) grew larger.

The number of missing cues varied across the experimental sessions and across diagnosticians depending upon the ease and coherence with which the picture of the reader was forming. (See Figure 9 for a comparative description of each client, and the hypotheses, diagnoses, and prescriptions offered by the diagnosticians in this first session.) Throughout the session, Sarah felt there was not enough information in the file, and the session was too short to really understand Tonita's reading problem. Judy's questions increased sharply after her interaction with Matt, to the point of not writing a diagnosis. She felt she did not have enough information to accurately identify his problem. Amy felt very confident with her interaction with Brett. She reported his problem as minor, and while wanting to have more time to work with him, felt her diagnosis and prescription was accurate. Across the three diagnostic sessions, as the number of missing cues and difficulty

of diagnosis increased, so did the number of inferences.

The majority of hypotheses were generated during the preparation stage, with each diagnostician stating that the interaction stage was the place to test tentative hypotheses. While Amy proposed a number of new hypotheses after the interaction, Judy and Sarah set forth most of their hypotheses early in the session. In fact, all three diagnosticians entertained their first hypothesis within the first four minutes of the preparation session. Given that this was the first experimental session of the study and the diagnosticians were relatively new at the think-aloud procedure and the context in general, this time may be overestimating the real time hypotheses begin to formulate. In later sessions, even in light of the differences in clients, hypotheses were offered many times within the first two minutes of the session.

In order to examine the consistency between hypotheses, diagnoses, and prescriptions within diagnosticians, a comparative analysis tracing the diagnoses back through the hypotheses and forward into the prescriptions was conducted. Figure 9 illustrates a brief picture of this comparison. In all cases, the entertained hypotheses, diagnoses, and prescriptions are related. In those areas that remediation may not be possible (e.g., Tonita's homelife), Sarah did not offer relevant diagnoses and prescriptions.

A look at the same client... Consistency? For experimental session V, Amy, Judy, and Sarah were asked to diagnose Peter's reading problem via a videotape of Peter working with another

diagnostician who was not a participant in this study. Peter is in third grade, has not repeated any grade, but was nearly retained the previous year. He is receiving help in the LD resource room and in Chapter I. The videotape interaction consisted of Peter completing a Peabody Picture Vocabulary Test, the Bryant Test of Phonics Skills, and the Spache Individual Reading Inventory.

The diagnosticians had up to 30 minutes to examine Peter's file before they watched the videotape. Prior to viewing the tape, they had to reveal the assessments they would administer in a real diagnostic interaction. They had one hour in which to view a one hour diagnostic session. Consequently, if they chose to stop and review parts of the tape, that would take up their hour.

Figure 10 outlines the order of cues selected during the preparation stage and after the interaction stage. All three diagnosticians appeared to engage in a purposive search through the file, citing time as the major impetus. As this session represented the second, third or fourth treatment session, the diagnosticians had become familiar with the scope of the file contents and were able to access salient information at a much faster and more efficient rate. Consequently, they were able to see more information within a short period of time, thereby increasing their knowledge about the upcoming client. Again, Amy was the only one to examine the entire file.

While Sarah stayed fairly close to the assessments she used with Tonita, she added a classroom observation in order to see Peter function in a structured environment. Sarah had noticed

information in the file that indicated part of his problem was related to his classroom work habits and socializing. Amy also wanted to conduct a classroom observation, along with one of her IRI's. In addition, she stated that math word problems would be helpful in order to test his comprehension in a preferred subject. She also desired an informal conversation with Peter to assess his motivation and attitude toward school and reading. Judy focused on the language difficulties she read about in the file, and decided that language assessments and the Slingerland Test for Learning Disabilities would be appropriate to corroborate the file's mention of his status in an LD resource room.

Table 2 reports the numbers of single cues, multiple cues, missing cues and inferences, and hypotheses. In her effort to examine the entire corpus of information on Peter, Amy spent more time (30 minutes) and consequently selected more single cues than Judy or Sarah who spent 24 and 20 minutes, respectively, in the preparation stage. Many of these cues came from the interaction, but Amy was also the only one to go back and access the remainder of the cues in the folder. Sarah and Judy selected approximately the same number of single cues, with Sarah verbalizing appreciably more during the stimulated recall stage than either Judy or Amy. Overall, all three selected more single cues in the beginning stages that toward the wrap-up.

Multiple cues were loaded toward the beginning of the diagnosis for Amy and Judy, but not Sarah -- hers were evenly distributed across stages. Again familiarity with the file

contents seemed to impact on this selection. The diagnosticians were grouping information such as grades, scores, and comments very quickly, and not taking the time to verbalize each available cue.

Many of the missing cues were filled with inferential statements. The missing cues for all diagnosticians focused the influence his apparent reading difficulty had on his classroom behavior. The videotape had captured the humorous side of Peter, with him making faces and hamming it up for the camera and the diagnostician. These displays of humor was interpreted by all three diagnosticians to mean that Peter was of average to above average intelligence and must be motivated to perform and cooperate in the classroom.

The hypotheses that led to many of these inferential statements were verbalized early in the experimental sessions. The diagnosticians had concluded early that his reading difficulties were phonetic, he was a strong visual learner, and he had a tendency to joke when frustrated or put in a difficult situation. (See Figure 11 for the comparative display of hypotheses, diagnoses, and prescriptions across diagnosticians.) All three indicated that Peter's teacher must be sensitive to his ability, interests, and interactional style. For each diagnostician, the diagnosis emanated from the selected cues and hypotheses, and were reflected in the prescriptions.

Implications and Scientific Importance

The most important implication of the current investigation may be that reading diagnosticians are not as internally

inconsistent as the Vinsonhaler research seemed to indicate. It appears likely that earlier studies may have actually created intra-diagnostician differences as artifacts of the artificial client simulations. Our findings indicate that expert diagnosticians rely on more kinds and complex interactions of cues and when information is not available, they will infer their best guess based on the evidence or withhold the diagnosis altogether. Given the restricted setting of the Vinsonhaler studies and the emphasis on the production of a written diagnosis for each clinical encounter, the participants may have had to make a number of inferences, resulting in the inconsistent diagnoses and prescriptions. We found that by tracing the entire problem identification process in more contextually valid situations, diagnoses along with the prescriptions were largely consistent. Further, we were able to examine a more complete picture of the diagnostic process by focusing on the process, beginning with the examination of single cues, building and grouping cues into multiple cue chunks, generating and testing hypotheses, and in most cases, concluding in the identification of the problem.

On the other hand, the data that indicated inter-diagnostician differences, similar to the Vinsonhaler studies, but different from Elstein and colleagues' medical diagnosis work, is perhaps neither surprising nor negative. As Kingsbury (1987) points out, medical schools tend to teach science as a set of facts rather than a method of inquiry. Given this orientation, differences in the training of physicians are

minimal. In addition, diagnosis is a particular emphasis of medical schools since diagnosis often implies treatment in rather direct fashion. Reading, however, is embedded in the social sciences with a wide range of theories and applications. Consequently, it would not be unlikely that reading diagnosticians would vary widely in orientation as a result of diverse training and experience.

Figure 1. A brief outline of the background experiences and diagnostic procedures and assessments for the three target diagnosticians.

Amy (reading specialist):

Background experiences: M.S. in Reading; 17 years teaching experience (4 years in the regular classroom, 13 years as a reading specialist).

Typical procedures: "I like to get impressions of kids. Usually I'll know the kids and the teachers. The first week of school I take the whole grade level's folders, flip through, looking for the SRA score, a pattern of scores (may have had one bad year). Then I like to do my own inventory, I like to make up my own mind then I would go to the files to look at the further testing. I collect information until something tends to be a picture. For instance, not noticing one score that seemed to stand out, not noticing the differences in the teacher's grades, but consistently all the way through. To some degree, I don't pay much attention to the previous reading, I do my inventory because I know the teacher and I might not evaluate the same."

Typical assessments: Teacher-made Individual Reading Inventory (IRI), IRI from basal series, or Silvaroli IRI (used with student from other schools, using other textbooks); teacher-made inventory of pre-reading and reading skills (e.g., phonetic and structural analysis); sight word list (e.g., Dolch); classroom observations

Judy (LD specialist):

Background experiences: M.S. in Elementary Education, specialization in learning disabilities; 7 years teaching experience with the learning disabled.

Typical procedures: "I tend to pull out the file and read it on Tuesday and test the child on Wednesday. I'm a real 'stew it over a while' person. I've always given myself a couple of hours. Normally I go through everything in their files and figure out what to do before ever selecting a test. I try to find a major weak area and match a test to that area. If there is any question in my mind, I don't want to stop. I always want to rule health problems out first. It's fun because when you're testing, it's like being a detective. Before every step, I formulate questions, then look for answers. The most important information is in working with the child, in the teaching situation. I usually do four hours of testing over a month, using informal and diagnostic teaching and classroom observations, kids can have bad days."

Typical assessments: Woodcock Reading Mastery Tests, Woodcock-Johnson Psycho-educational Battery of Tests, Key Math, Slingerland Test for Learning Disabilities, Peabody Picture Vocabulary Test, informal (diagnostic) teaching sessions, miscue analyses, teachers', supervisor's, and psychologist's input, classroom observations.

Figure 1. (cont.). A brief outline of the background experiences and diagnostic procedures and assessments for the three target diagnosticians.

Sarah (classroom teacher):

Background experiences: M.S. in Elementary Education; 17 years experience in the regular classroom.

Typical procedures: "I don't put a lot of worth on past tests, the day-to-day interaction with the child is what I tune into a lot more. The one-on-one, I feel I get a lot more out of that, just put me with a child and let me talk with her for 10 minutes. I like to have a picture of the child to know what they might want to read. I want to see how she does on interacting with me on the stories. I like to hear them read. I also do the Dolch words and blends/cluster flash cards. An hour is too long, I usually work in smaller periods of time (5-15 minutes). I do a lot of observing and asking. I get ideas from the Chapter I teacher."

Typical assessments: Dolch sight words, flash cards, phonics kit, oral reading from basal reader or trade book (most often reading is from library books chosen by the child and/or the teacher), guided questioning after silent reading, classroom observations.

Figure 2. Description of the five experimental sessions.

Session #	Nature of client and diagnostic interaction	Order of treatment tasks *
I.	Live client: each diagnostician works with a different child.	Preparation, interaction, written diagnosis, stimulated recall, summary
II.	"Boxed" client: replication of Vinsonhaler studies; each diagnostician reviews the file and audiotapes of a different child.	Preparation (includes "interaction" with audiotaped reading), written diagnosis, stimulated recall, summary
III.	Live client: each diagnostician works with a different child (also different child than diagnosed in session 1).	Preparation, interaction, stimulated recall, written diagnosis, summary
IV.	Videotaped client: using the session 1 videotapes, each diagnostician identifies the reading difficulty(ies) of clients working with other diagnosticians in this study.	Preparation, interaction (both of these stages with think-alouds), written diagnosis, summary
V.	Videotaped client: control session; all diagnosticians identify the reading difficulties of a child who is working with a diagnostician not in this study.	Preparation, interaction, written diagnosis, stimulated recall, summary

* Descriptive think-alouds were used in all preparation and interaction stages, except where noted in experimental session IV.

Figure 3. The Case Information Inventory master list of cues available to the diagnosticians for each experimental session (i.e., availability dependant upon individual client's actual file).

I. General Information

- I-A. Referral to the Reading Clinic (comments from the school)
- I-B. Referral for school-related special services
- I-C. General/family background

II. Health/medical information

- II-A. History of childhood diseases
- II-B. Immunization history
- II-C. Height/weight records
- II-D. Vision
- II-E. Hearing
- II-F. Teeth
- II-G. Throat
- II-H. Speech

III. School Information

- III-A. Attendance
- III-B. Academic Progress (report cards)

IV. Tests/assessments administered by the School

- IV-A. Kindergarten screening
- IV-A*. Kindergarten Skills Checklist (for only one client)
- IV-B. Metropolitan Readiness Test
- IV-C. Curriculum Referenced Tests of Mastery
- IV-D. SRA Skills Assessment
- IV-E. Open Court Headway / Virginia Standards of Learning

IV. Tests/ Assessments administered by the School (Cont.)

- IV-F. Woodcock-Johnson Psychoeducational Battery of Tests
- IV-G. Weschler Intelligence Scale for Children - Revised
- IV-H. Bender Gestalt Motor Test
- IV-I. Slingerland Screening for Specific Language Disability
- IV-J. Test of Written Language
- IV-K. The Visual Aural Digit Span Test
- IV-L. Kaufman Assessment Battery for Children
- IV-M. Adaptive Behavior Evaluation Scale
- IV-N. Peabody Picture Vocabulary Test

V. Tests/information reported by the Reading Clinic

- V-A. Tutor's Report
- V-B. Informal Reading Inventory
- V-C. Peabody Picture Vocabulary Test
- V-D. Bryant Diagnostic Test of Phonics Skills
- V-E. Woodcock Reading Mastery Tests

VI. Tests/information reported by "Outside Reading Diagnostician"

- VI-A. Peabody Picture Vocabulary Test
- VI-B. Spache Informal Reading Inventory
- VI-C. Bryant Test of Phonics Skills
- VI-D. Informal written language assessment

Figure 4. Sequence with which the five diagnosticians interacted with the six clients.*

Diagnostician	1	2	3	4	5
Amy	(Brett)	Tonita	Matt	[Peter]	Stephanie
Judy	(Matt)	Barbara	[Peter]	Brett	Tonita
Pat	Stephanie	Barbara	Paul	Matt	Tonita
Bonnie	Barbara	Stephanie	Matt	Brett	Paul
Sarah	(Tonita)	[Peter]	Stephanie	Barbara	Brett

○ Comparison of the processes of the three target diagnosticians with three live clients in naturalistic conditions.

□ Comparison of the processes of the three target diagnosticians given the same experimental session with the same client (i.e., control condition).

Figure 5. Sequence of the experimental sessions across diagnosticians.

Diagnostician	Experimental Session Order				
	1	2	3	4	5
Amy	(I)	II	IV	[V]	III
Pat	I	II	V	III	IV
Judy	(I)	IV	[V]	II	III
Sarah	(I)	[V]	II	III	IV
Bonnie	I	IV	II	III	V

(I) Experimental session used to compare the processes of the three target diagnosticians with three different live clients.

[V] Comparison of the processes of the three target diagnosticians given the same experimental session with the same client (i.e., control condition).

Figure 6. The taxonomy of single cues requested by the three diagnosticians across the experimental sessions.

General/Family Background

From the records:

- Age
- Grade
- Birthdate
- Birthplace
- Father's education
- Mother's education
- No. of brothers, ages
- No. of sisters, ages
- Family status
- Misc. information (e.g., rides the bus)

From interaction:

- Interests
- Favorite toys, games, books
- Bedtime
- Television watching habits
- Time spent on homework or pleasure reading, or play
- Parents' speech patterns
- Living arrangements and visitation habits with divorced parents
- Responsibilities at home

Health/Medical Information

From the records:

- Childhood diseases
- Immunization
- Height/weight
- Vision
- Hearing
- Teeth
- Throat
- Speech

From interaction:

- Physical characteristics
- Cleanliness
- Neatness
- Clothing
- Physical well-being
- Consistency of wearing prescribed glasses

School Information

From the records:

- Academic progress
  - -grading system
  - -grades
  - -effort
  - -textbook series
  - -book/grade level
  - -teacher comments
    - academic
    - behavioral
    - suggestions for improvement
    - encouragement/ warnings
  - -retention history
- Attendance
- Special services
  - -referrals
  - -dates of referrals
  - -teacher comments
    - academic
    - behavioral
    - tentative hypotheses
  - -present placements
  - -parent involvement
  - -history (years in programs)

From interaction/general knowledge:

- Information about the school division
- Community socio-economic status
- Comparison of client to classmates
- Expectations of students in school division or by the teacher
- Availability of programs and special assistance
- General attitude toward school, school subjects, teachers
- Size of reading group

General Intellectual Ability

From records/stand. tests:

- Name of test
- Nature of test
- Grade equivalent
- Standard score
- Percentile
- Raw scores
- Errors
- Subtest scores
- Examiner's interpretations
- Examiner's notes about client's behavior during test
- Examiner's diagnosis
- Examiner's recommendations

From interaction:

- Ease, clarity, and depth of conversation
- Reactions to questions
- Reaction time to questions
- No. of prompts used
- Kinds of prompts needed
- No. of times questions had to be restated
- Speech patterns
- Vocabulary
- Use of details and explanations
- Depth of prior knowledge
- Humor
- Ability to follow a complex set of directions

Figure 6 (cont.). The taxonomy of single cues requested by the three diagnosticians across experimental sessions

Level of Achievement

From standardized tests on record:

- Criterion referenced tests
  - -grade equivalent
  - -standard scores
  - -percentiles
  - -skills mastered/not mastered
  - -raw score
- Norm-referenced tests
  - -grade equivalent
  - -standard scores
  - -percentiles
  - -subtest scores
  - -miscues

From tutor's report:

- No. of books read, examples
- Sight word performance
- No. of stories written
- Attitude
- General comments
- Recommendations

From informal inventories (on record or administered during session:

- Word recognition list
  - -miscues related to phonetic and structural analysis
  - -reading level
  - -order words are read
  - -raw score
  - -percent correct
- Comprehension passages
  - -reading level
  - -miscues
  - -use of context
  - -corrections
  - -if errors made sense
  - -responses to comp. questions
    - type missed
    - depth of answer
    - completeness
    - speech patterns
    - speed of response

- Oral reading behavior
  - -effort
    - desire to continue task
    - attempt to attack words
    - attempt to answer ques.
  - -fluency
    - ran out of breath
    - swallows
    - phrasing, word-by-word reading
    - hesitations, stops
    - pauses
  - -expression
    - voice
    - intonations
    - attention to punctuation
  - -humming during silent reading
  - -sighs
  - -escape behavior when asked to respond
  - -use of finger, etc. as placeholder

Other Cues from the Testing Situation:

Contextual cues:

- Lighting
- Table/chair height
- Time of day
- Time to gather information
- Availability of information
- Time to test
- Nature of tests (e.g., nonsense vs. real words)
- Testing materials (e.g., scoring)
- Examiner (in audio/video sessions)
  - -background
  - -present assignment
  - -interaction w/ client

Body language/non-verbal cues:

- Attention to the situation (e.g., looking around the room)
- Listening
- Proximity to testing materials and/or diagnostician
- Head movement during reading
- Order pictures are examined
- Use of fingers for reading
- Looking at diagnostician, pausing for help
- Restlessness
- Visible signs of stress (e.g., facial contortions)
- Yawning, stretching
- "Fiddling" with glasses

Figure 7. The taxonomy of hypotheses considered by the three diagnosticians across the experimental sessions.

General/Family Background

- Quality/amount of intellectual stimulation in home environment
- Stability of family (two vs. single parent home, divorces/remarried parents)
- Number of siblings, birth order
- Parental expectations, overprotective nature, willingness to accept child's problems

School Context

- Compatibility with the teacher
- Compatibility with the reading series
- Change of school context (e.g., due to moving, change in classroom structure)
- Expectations of the teacher, child's classmates, school, community
- Motivation to learn in school

Reading Skills

- Phonics
- Comprehension (e.g., literal, inferential)
- "Page turner" -- poor oral reading habits
- Sight vocabulary

Health/Medical/Organic Problems

- Lasting impact of childhood diseases (e.g., high fever) or injuries (e.g., concussion)
- Allergies
- Hearing, Vision and/or Speech deficits
- History of medications, if on medication presently for allergies, hyperactivity)
- Large/small muscle coordination
- Immaturity, delayed development
- Neurological disorders
- Attendance/illnesses
- Perceptual disorders
- Attentional disorder/Fatigue level-tolerance

Intellectual Ability/ Ability to Learn

- Ability level/potential
- General knowledge/experiences
- Learning disability
- Memory deficit/disorder (i.e., recall, recognition, reconstruction)
- Metacognitive abilities
- Learning modality
- Language disorder

General

- Self-concept
- Ability to handle frustration
- Interaction among any of the classes of cues/hypotheses

Figure 8. The order of cues selected from the Case Information Inventory (cumulative file) by the three target diagnosticians before the interaction, the assessments used during the interaction, and the subsequent cues requested prior to the diagnosis in experimental session 1.

Diagnostician	Before working with the client	Assessments used	Additional cues requested
Amy *	Proceeded in order of cues on Case Inventory. Selected clinic referral, family background, childhood diseases, vision, hearing, speech, attendance, kindergarten report card, kindergarten skills checklist, 1st grade report card, all reports of SRA test results, IRI from the clinic, IRI from outside diagnostician	Basal Informal Reading Inventory; informal inventory of readiness and reading skills (consonants, consonant blends, digraphs, 3-letter clusters, silent letters, short and long vowel sounds)	Examined previously unseen SRA subtest results.
Judy	Selected cues in purposive manner. Woodcock-Johnson Psychoeducational Battery of Tests, WISC-R, Bender-Gestalt Test, Slingerland Screening for LD, attendance, vision, hearing, speech, Kindergarten report card (all subjects), Kindergarten-repeat report card (rdg., math, teacher comments), first grade report card (all subjects), Metropolitan Readiness Test, remainder of family background, IRI from clinic, Test of Written Language	Informal written lang. assessment, Peabody Picture Vocabulary Test, Woodcock Reading Mastery Tests	Examined previously unseen subtests and comments on WISC-R, clinic's tutor's report
Sarah	Selected cues in purposive manner. Family background, referral to clinic, vision, hearing, attendance, second grade report card (work habits and core curricular subjects, last marking period's teacher comments), third grade report card (behavior, work habits, core curricular subjects, all teacher comments, 2nd grade Criterion Referenced Test of Mastery (CRTM), 3rd grade CRTM (only those areas not mastered previous year), tutor's report from clinic, tutor's IRI, Bryant Test of Phonics	Informal reading of books chosen by client and diag. together; Ekwall IRI.	None

* Examined entire file.

Table 1. Amounts of different cues and hypotheses requested/ reported by the three target diagnosticians in each stage of experimental session 1 (interaction with a live client).*

	NUMBER OF DIFFERENT CUES AND HYPOTHESES VERBALIZED AT EACH STAGE									TOTAL VERBAL-IZED
AMY										
	Prep.	N./Prep.	N. Int.	Add. Cues	Notes/AC	Diag.	Stim. Recall	Wrap	Total	
Single Cues	65	32	113	0	12	8	29	0	259	421
Multiple Cues	23	7	0	0	0	31	23	5	89	105
Missing Cues	15	0	0	0	0	1	3	2	21	23
Inferences	2	0	0	0	0	0	0	0	2	2
Hypotheses	7	1	0	0	0	3	9	3	23	43
JUDY										
	Prep.	N./Prep.	N. Int.	Add. Cues	Notes/AC	Diag.	Stim. Recall	Wrap	Total	
Single Cues	69	5	82	7	0	10	63	0	236	428
Multiple Cues	47	0	0	0	0	11	47	0	105	145
Missing Cues	10	0	0	0	0	0	24	1	35	51
Inferences	2	0	0	0	0	0	17	0	19	19
Hypotheses	8	0	0	0	0	0	5	0	13	123
SARAH										
	Prep.	N./Prep.	N. Int.	Add. Cues	Notes/AC	Diag.	Stim. Recall	Wrap	Total	
Single Cues	73	1	113	0	0	13	85	1	286	507
Multiple Cues	38	0	6	0	0	7	55	3	109	151
Missing Cues	33	2	0	0	0	0	33	1	69	86
Inferences	10	1	0	0	0	0	16	1	28	28
Hypotheses	7	0	0	0	0	0	0	0	10	158

* Both Amy and Judy wanted more time to assess their clients, while Sarah did not. All three diagnosticians reached a tentative diagnosis.

Figure 9. A comparative display of the major hypotheses, diagnoses and prescriptions across diagnosticians in experimental session 1.

	Amy	Judy	Sasah
Client	Brett: 2nd grader, Chapter 1, recommended for LD testing, parent refuse	Matt: 2nd grader, repeated kindergarten, is reported to have aud. discrim problems (not phonetic), having some trouble staying out of trouble. Teacher observing math must better than reading/language.	Tonita: 3rd grader, repeating this year, in Chapter 1, no referral from school, no detail in clinic referral.
Hypotheses	Parent high expectations. Auditory/phonics weaker area. Mismatch between basal program and student. Hesitancy/choppiness during oral reading is greatest problem.	Learning disability in language. Behavior problem due to low self-concept	Homelife: divorce, both parents gone, lives with granny. Probable cause of drastic drop in third grade (eventual retention) "Page turner" -- reads too fast, misses details, impedes comprehension
Diagnoses	Instructional level approx 2.1, comprehension is good, not severely disabled, average ability. Good verbal expression. Oral reading fluency most evident problem.	His behavior is not a good indicator of his ability. Has some major problem related to putting words together (tests don't yet show it). Still have many questions, recommend further testing. No diagnosis yet.	Doing well despite home life Problem with literal comp. fast reader, ignores punctuation. About on grade level for reading, sounds on configuration, needs some brushing up on phonics.
Prescription	Lots of easy reading to practice phrasing, work with phrasing cards. Try a different basal. Needs a review and practice on long & short vowels.	LD resource services by the speech/language therapist Discontinue clinic program, reward him for progressing and doing well.	Have her tape her reading and play it back. Make her aware of how she sounds. Work on speed needs to hear models. Do assisted reading. Focus on comprehension checks of basic facts.

Figure 10. The order of cues selected from Peter's Case Information Inventory (cumulative file) by the three target diagnosticians before the interaction, the preferred assessments during the interaction and the subsequent cues requested prior to the diagnosis in experimental session 5.

Diagnostician	Before observing the client on videotape	Assessments would have used	Additional cues requested
Amy *	Selected cues in a purposive manner. Referral to the clinic; kindergarten screeing; Metropolitan Readiness Test; 1st and 2nd grade CRTM; attendance; kindergarten, first, and second grade report cards; referral for school services; family background	Would have started w/ Silvaroli IRI, then classroom observation; informal conversation with Peter; math word problems	All the psychological testing (Woodcock Johnson, WISC-R, Bender-Gestalt, Visual Aural Digit Span Test); all the clinic reports; remainder of file on health/medical
Judy	Selected cues in a purposive manner. Referral to the clinic; vision, hearing, speech; attendance; kindergarten report card; 1st grade report card; second grade report card; kindergarten screening (lang. only); WISC-R; Woodcock Johnson Psychoeducational Battery; clinic IRI; referral for school services; family background; Metropolitan Readiness Test; 1st and 2nd grade CRTM; Bender-Gestalt Test; Visual Aural Digit Span Test	Slingerland Test for LD; list of short vowel sounds to sound out; a language development assessment (e.g., TOLD); a language sample	None
Sarah	Selected cues in purposive manner. Referral to the clinic; referral for school special services; parent occupations and comments about family; vision, hearing; attendance, kindergarten lang. development grade; first grade core subjects and last marking period teacher's comment; 2nd grade work habits, rdg., lang., spelling; 1st and 2nd grade CRTM; Woodcock Johnson Psychoeducational Battery; WISC-R; Tutor's report; clinic IRI; Visual Aural Digit Span Test	Informal reading of books chosen by client and diag. together; Ekwall IRI; classroom observation.	Second grade teacher comments for entire year.

* Examined the entire file.

Table 2. Amounts of different cues and hypotheses requested/ reported by the three target diagnosticians in each stage of experimental session 5 (same client under same conditions).*

	NUMBER OF DIFFERENT CUES AND HYPOTHESES VERBALIZED AT EACH STAGE										TOTAL VERBAL-IZED
AMY											
	Prep.	N./Prep.	Int.	N. Int.	Add. Cues	Notes/AC	Diag.	Stim. Recall	Wrap	Total	
Single Cues	115	0	45	134	7	28	1	24	0	348	516
Multiple Cues	34	0	8	6	0	0	2	22	2	74	103
Missing Cues	19	0	8	1	0	0	0	14	1	43	54
Inferences	8	0	7	0	0	0	0	6	1	22	22
Hypotheses	7	0	3	1	0	0	0	3	0	14	45
JUDY											
	Prep.	N./Prep.	Int.	N. Int.	Add. Cues	Notes/AC	Diag.	Stim. Recall	Wrap	Total	
Single Cues	82	2	82	22	0	0	4	11	0	203	270
Multiple Cues	51	0	24	1	0	0	2	16	0	94	158
Missing Cues	13	0	23	2	0	0	1	0	0	44	67
Inferences	5	0	23	2	0	0	0	0	0	28	28
Hypotheses	5	0	6	0	0	0	0	0	0	11	42
SARAH											
	Prep.	N./Prep.	Int.	N. Int.	Add. Cues	Notes/AC	Diag.	Stim. Recall	Wrap	Total	
Single Cues	87	2	94	0	0	0	10	16	4	213	427
Multiple Cues	35	29	1	0	0	0	4	30	3	102	147
Missing Cues	10	0	26	2	0	0	4	9	0	51	67
Inferences	4	0	16	0	0	0	4	5	0	29	29
Hypotheses	13	0	3	0	0	0	0	0	0	16	167

* Both Amy and Judy wanted more time to assess their clients, while Sarah did not. All three diagnosticians reached a tentative diagnosis.

Figure 11. A comparative display of the major hypotheses, diagnoses and prescriptions across diagnosticians in experimental session 5

	Amy	Judy	Sarah
Hypotheses	Phonetic problems, relates to WISC-R object assembly. Visual learner, relies on context does not pay attention to details. Would be difficult to keep his attention in a larger group.	Language/vocabulary difficulties, does not seem to retain vocab. Poor in phonics, seems to have auditory discrimination problem. Body language indicates frustration	Phonics problem, doesn't really know the rules and how to apply them. Must be motivated to attend, hold interest; attention span appears short Jokes, goes into body contortions when frustrated Depends upon context for meaning/decoding
Diagnoses	Average ability; reading level is about 3.0 (might want to test further) Appears to think, get meaning Uses context; weak in word attack, syllabication	Needs Slingerland to check on other perceptual problems. Not firm on LD diagnoses, but problem is severe enough cause learning problems. Has a good sense of humor. Average student. Problems integrating parts to whole	Should be performing above what he is presently doing. Good sense of humor, creative. Body language gives away his attitude. Has fallen a little behind in skills. Can't do phonics, should stay away.
Prescription	Definitely needs to work on vowels; teach vowels in real, whole words. Must motivate him, use high interest informal materials.	Give clear intro to words. Be aware of body language as a sign of frustration. Tell him how to chunk letters & words together.	Needs high interest material. Use multisensory materials. Needs to be kept busy, work for short periods of time. Give individual guidance on word attack.

References

Adelson, B. (1981). Problem solving and the development of abstract categories in programming languages. Memory and Cognition, 9, 422-433.

Berliner, D.C. (1986). In pursuit of the expert pedagogue. Educational Researcher, 15 (7), 5-13.

Bloom, B.S. (1986). Automaticity: "The hands and feet of genius." Educational Leadership, 83 (5), 70-77.

Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). A study of thinking. New York: John Wiley & Sons.

Chase, W.G., & Ericsson, K.A. (1981). Skilled memory. In J.R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

DeGroot, A.D. (1966). Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmuntz (Ed.), Problem solving. New York: John Wiley & Sons.

Elstein, A.S., Kagan, N., Shulman, L.S., Jason, P.H., & Loupe, M.J. (1972). Methods and theory in the study of medical inquiry. Journal of Medical Education, 47, 85-92.

Elstein, A.S., Shulman, L.S., & Sprafka, S. (1978a). Medical problem solving: An analysis of clinical

reasoning . Cambridge: Harvard University Press.

Ericcson, K.A., & Simon, H.A. (1984). Protocol analysis : Verbal reports as data . Cambridge, MA: The MIT Press.

Gil, D., Hoffmeyer, E., VanRoekel, J., Vinsonhaler, J., & Weinshank, A. (1979). Clinical problem solving in reading: Theory and research (Research Series No. 45). East Lansing, MI: Michigan State University, Institute for Research on Teaching.

Hammond, K.R., & Summers, D.A. (1972). Cognitive control. Psychological Review , 79 , 58-67.

Housner, L.D. & Griffey, D.C. (1985). Teacher cognition: Differences in planning and interactive decision-making between experienced and inexperienced teachers. Research Quarterly for Exercise and Sport, 56, 45-53.

Lee, A., & Weinshank, A. (1978). Case production and analysis - CLIPIR pilot observational study, 1976. (Research Series No. 14). East Lansing, MI: Michigan State University, Institute for Research on Teaching.

Leinhardt, G., & Greeno, J.G. (1986). The cognitive skill of teaching. Journal of Educational Psychology , 78 , 75-95.

Neves, D.M., & Anderson, J.R. (1981). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J.R. Anderson (Ed.), Cognitive skills and their acquisition . Hillsdale, NJ: Lawrence Erlbaum Associates.

Newell, A., & Simon, H.A. (1972). Human problem solving . Englewood Cliffs, NJ: Prentice-Hall, Inc.

Polya, G. (1957). How to solve it . Garden City, NY: Doubleday Anchor.

Polya, G. (1968). Mathematical discovery. Vol. II: On understanding, learning and teaching problem solving . New York: John Wiley & Sons.

Schon, D.A. (1983). The reflective practitioner: How professionals think in action . New York: Basic Books.

Seibel, R. (1963). Discrimination reaction time for a 1,023 alternative task. Journal of Experimental Psychology , 66 , 215-226.

Shulman, L.S., & Elstein, A.S. (1975). Studies of problem solving, judgment, and decision making: Implication for educational research. In F.N. Kerlinger (Ed.), Review of research in education (Vol. 3) . Itasca, IL: F.E. Peacock.

Slovic, P., & MacPhillamy, D.J. (1974). Dimensional commensurability and cue utilization in comparative judgment. Organizational Behavior and Human Performance , 11 , 172-194.

Spradley, J.P. (1980). Participant observation . New York: Holt, Rinehart & Winston.

Trabasso, T.R. (1963). Stimulus emphasis and all-or-none learning in concept identification. Journal of Experimental Psychology , 65 , 398-399.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory.

Psychological Review , 80 , 352-373.

Vinsonhaler, J.F., Weinshank, A.B., Wagner, C.C., & Polin, R.M. (1983). Diagnosing children with educational problems: Characteristics of reading and learning disabilities specialists, and classroom teachers. Reading Research Quarterly , 18 , 134-164.